

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: O'Malley, Bert W.  
Tsai, Ming-Jer  
Ledebur, Harry C. Jr.  
Kittle, Joseph D. Jr.
- (ii) TITLE OF INVENTION: MODIFIED STEROID  
HORMONES FOR GENE  
THERAPY AND METHODS  
FOR THEIR USE
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Lyon & Lyon  
(B) STREET: 633 West Fifth Street  
Suite 4700  
(C) CITY: Los Angeles  
(D) STATE: California  
(E) COUNTRY: U.S.A.  
(F) ZIP: 90071-2066
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: IBM P.C. DOS 5.0  
(D) SOFTWARE: Word Perfect 5.1
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: To Be Assigned  
(B) FILING DATE: Herewith  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: 08/479,913  
(B) FILING DATE: June 7, 1995
- (A) APPLICATION NUMBER: 07/939,246  
(B) FILING DATE: September 2, 1992

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warburg, Richard J.  
 (B) REGISTRATION NUMBER: 32,327  
 (C) REFERENCE/DOCKET NUMBER: 222/085

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (213) 489-1600  
 (B) TELEFAX: (213) 955-0440  
 (C) TELEX: 67-3510

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6177 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: nucleic acid

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTAGAGTCGA	CCTGCAGCCC	AAGCTCTCGA	GGGATCCTGA	GAACCTCAGG	GTGAGTTTGG	60
GGACCCTTGA	TTGTTCTTTC	TTTTTCGCTA	TTGTAAAT	CATGTTATAT	GGAGGGGGCA	120
AAGTTTTTCAG	GGTGTGTGTT	AGAATGGGAA	GATGTCCCTT	GTATCACCAT	GGACCCTCAT	180
GATAAATTTTG	TTTCTTTTAC	TTTCTACTCT	GTTGACAACC	ATTGTCTCCT	CTTATTTTCT	240
TTTCATTTTTC	TGTAACCTTT	TCGTAAACT	TTAGCTTGCA	TTTGTAACGA	ATTTTTAAAT	300
TCACCTTTTGT	TTATTTGTCA	GATTGTAACT	ACTTTCTCTA	ATCACTTTT	TTTCAAGGCA	360
ATCAGGGTAT	ATTATATTGT	ACTTCAGCAC	AGTTTTAGAG	AACAATTGTT	ATAATTAAAT	420
GATAAGGTAG	AATATTTCTG	CATATAAAT	CTGGCTGGCG	TGGAAATATT	CTTATTGGTA	480
GAAACAACATA	CATCCTGGTC	ATCATCTGTC	CTTTCTCTTT	ATGGTTACAA	TGATATACAC	540
TGTTTGAGAT	GAGGATAAAA	TACTCTGAGT	CCAAACCGGG	CCCCTCTGCT	AACCATGTTT	600
ATGCCTTCTT	CTTTTTCCCTA	CAGCTCCTGG	GCAACGTGCT	GGTTGTGTG	CTGTCTCATC	660
ATTTTGGCAA	AGAATTCAC	CCTCAGGTGC	AGGCTGCCTA	TCAGAAGGTG	GTGGCTGGTG	720
TGGCCAATGC	CCTGGCTCAC	AAATACCACT	GAGATCTTTT	TCCCTCTGCC	AAAAATTATG	780
GGGACATCAT	GAGCCCCCTT	GAGCATCTGA	CTTCTGGCTA	ATAAAGGAAA	TTTATTTTCA	840
TTGCAATAGT	GTGTTGGAAT	TTTTTGTGTC	TCTCACTCGG	AAGGACATAT	GGGAGGGCAA	900
ATCATTTAAA	ACATCAGAA	GAGTATTTGG	TTTAGAGTTT	GGCAACATAT	GCCATATGCT	960
GGCTGCCATG	AACAAAGGTG	GCTATAAAGA	GGTCATCAGT	ATATGAAACA	GCCCCCTGCT	1020
TGCCATTCCCT	TATTCATAG	AAAAGCCTTG	ACTTGAGGTT	AGATTTTTT	TATATTTTGT	1080
TTTGTGTTAT	TTTTTCTTT	AACATCCCTA	AAATTTTCTT	TACATGTTT	ACTAGCCAGA	1140
TTTTTCCTCC	TCTCCTGACT	ACTCCCAGTC	ATAGCTGTCC	CTCTTCTCTT	ATGAACTCGA	1200
GGAGCTTTTT	GCAAAAGCCT	AGGCCTCCAA	AAAAGCCTCC	TCACTACTTC	TGGAATAGCT	1260
CAGAGGCCGA	GGCGGCCCTG	GCCTCTGCAT	AAATAAAAAA	AATTAGTCAG	CCATGGGGCG	1320
GAGAATGGGC	GGAAGTGGGC	GGAGTTAGGG	GCGGGATGGG	CGGAGTTAGG	GGCGGGACTA	1380
TGGTTGTGTA	CTAATTGAGA	CTGCATTAAT	GAATCGGCCA	ACGCGCGGGG	AGAGGCGGTT	1440
TGCGTATTGG	GCGCTCTTCC	GCTTCTCGC	TCTACTGACT	GCTGCGCTCG	GTCTGTCGGC	1500
TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAATACG	GTTATCCACA	GAATCAGGGG	1560
ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG	1620
CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGTA	CGAGCATCAC	AAAAATCGAC	1680
GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	1740
GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	1800
TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC	AATGCTCAG	CTGTAGGTAT	CTCAGTTCCG	1860
TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCA	CCCGACCGCT	1920
GCGCTTTATC	CGGTAACAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	TTATCGCCAC	1980
TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	2040
TCTTGAAGTG	TGGGCCTAAC	TACGGCTACA	CTAGAAGGAC	AGTATTTGGT	ATCTGCGCTC	2100
TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC	AAACAAACCA	2160
CCGCTGGTAG	CGGTGGTTTT	TTTGTTTTGA	AGCAGCAGAT	TACGCGCAGA	AAAAAAGGAT	2220

CTCAAGAAGA	TCCTTTGATC	TTTTCTACGG	GGTCTGACGC	TCAGTGGAAC	GAAAACTCAC	2280
GTTAAGGGAT	TTTGGTTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC	CTTTTAAAT	2340
AAAAATGAAG	TTTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTTGGTCT	GACAGTTACC	2400
AATGCTTAAT	CAGTGAGGCA	CCTATCTCAG	CGATCTGTCT	ATTTTCGTTCA	TCCATAGTTG	2460
CCTGACTCCC	CGTCGTGTAG	ATAACTACGA	TACGGGAGGG	CTTACCATCT	GGCCCCAGTG	2520
CTGCAATGAT	ACCGCGAGAC	CCACGCTCAC	CGGCTCCAGA	TTTATCAGCA	ATAAACCCAGC	2580
CAGCCGGAAG	GGCCGAGCGC	AGAAGTGGTC	CTGCAACTTT	ATCCGCCTCC	ATCCAGTCTA	2640
TTAATTGTTG	CCGGGAAGCT	AGAGTAAGTA	GTTCCGCCAGT	TAATAGTTTG	CGCAACGTTG	2700
TTGCCATTGC	TACAGGCATC	GTGGTGTAC	GCTCGTCGTT	TGGTATGGCT	TCATTCCAGCT	2760
CCGGTTCCCC	ACGATCAAGG	CGAGTTACAT	GATCCCCCAT	GTTGTGCAAA	AAAGCGGTTA	2820
GCTCCTTCGG	TCCGTCGATC	GTTGTGAGAA	GTAAGTTGGC	CGCAGTGTTA	TCACTCATGG	2880
TTATGGCAGC	ACTGCATAAT	TCTCTTACTG	TCATGCCATC	CGTAAGATGC	TTTTCTGTGA	2940
CTGGTGAGTA	CTCAACCAAG	TCATTCTGAG	AATAGTGTAT	GCGGCGACCG	AGTTGCTCTT	3000
GCCCGGCGTC	AATACGGGAT	AATACCGCGC	CACATAGCAG	AACTTTAAAA	GTGCTCATCA	3060
TTGGAACACG	TTCTTCGGGG	CGAAAACTCT	CAAGGATCTT	ACCGCTGTTG	AGATCCAGTT	3120
CGATGTAACC	CACCTCGTGCA	CCCAACTGAT	CTTCAGCATC	TTTTACTTTC	ACCAGCGTTT	3180
CTGGGTGAGC	AAAAACAGGA	AGGCAAAATG	CCGCAAAAAA	GGAATAAGG	GCGACACGGA	3240
AATGTTGAAT	ACTCATACTC	TTCTTTTTC	AATATTATTG	AAGCATTTAT	CAGGGTTATT	3300
GTCTCATGAG	CGGATACATA	TTTGAATGTA	TTTAGAAAAA	TAAACAAATA	GGGGTTCCGC	3360
GCACATTTCC	CCGAAAAGTG	CCACCTGACG	TCTAAGAAAC	CATTATTATC	ATGACATTAA	3420
CCTATAAAAA	TAGGCGTATC	ACGAGGCCCT	TTCTCTTCA	AGCTGCCCTC	CGCGTTTCGG	3480
TGATGACGGT	GAAAACTCT	GACACATGCA	GCTCCCGGAG	ACGGTCACAG	CTTGTCTGTA	3540
AGCGGATGCC	GGGAGCAGAC	AAGCCCGTCA	GGGCGCGTCA	GCGGGTGTG	GCGGGTGTG	3600
GGCGCAGCC	ATGACCCAGT	CACGTAGCGA	TAGCGGAGTT	GGCTTAAC	TGCGGCATCA	3660
GAGCAGATTG	TACTGAGAGT	GCACCATATC	GACGCTCTCC	CTTATGCGAC	TCCTGCATTA	3720
GGAAGCAGCC	CAGTAGTAGG	TTGAGGCCGT	TGAGCACC	CGCCGCAAGG	AATGGTGTCTG	3780
GCTTATCGAA	ATTAATCGAC	TCACTATAGG	GAGACCCGAA	TTTCGAGCTC	CCCCGTTACA	3840
TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCCAAG	ACCCCGCCCC	ATTGACGTCA	3900
ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	TCAATGGGTG	3960
GAGTATTTAC	GGTAAACTGC	CCACTTGGCA	GTACATCAAG	TGTATCATAT	GCCAAGTACG	4020
CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCCCTCTGGC	ATTATGCCCA	GTACATGACC	4080
TTATGGGACT	TTCTTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGTG	4140
ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	GGGATTTCCA	4200
AGTCTCCACC	CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	ACGGGACTTT	4260
CCAAAAATGC	GTAACAAC	CGCCCCATTG	ACGCAAAATG	GCGGTAGGCG	TGTACGGTGG	4320
GAGGTCTATA	TAAAGCAGAG	TCGTTTAGTG	AACCGTCAGA	TGCGCTGGAG	ACGCCATCCA	4380
CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	GATCTTGGTG	4440
GCGTGAACT	CCCGCACCTC	TTCCGGCCAGC	GCCTTGTAGA	AGCGCGTATG	GCTTCGTGGG	4500
GATCCCCCAA	AGAACTCTTA	GCTCCCCCTG	GTAGAGACGA	AGTCCCTGGC	AGTTTGGCTG	4560
GCCAAGGGAG	GGGGAGCGTA	ATGGACTTTT	ATAAAAGCCT	GAGGGGAGGA	GCTACAGTCA	4620
AGGTTTCTGC	ATCTTCGCCC	TCAGTGGCTG	CTGCTTCTCA	GGCAGATTCC	AAGCAGCAGA	4680
GGATTCTCCT	TGATTTCTCG	AAAGGCTCCA	CAAGCAATGT	GCAGCAGCGA	CAGCAGCAGC	4740
AGCAGCAGCA	GACAGCAGAG	CAGCAGCAGC	AGCAGCAGCA	GCAGCAGCCA	GGCTTATCCA	4800
AAGCCGTTTC	ACTGTCCATG	GGGCTGTATA	TGGGAGAGAC	AGAAACAAAA	GTGATGGGGA	4860
ATGACTTGGG	CTACCCACAG	CAGGGCCAAC	TTGGCCTTTC	CTCTGGGGAA	ACAGACTTTC	4920
GGCTTCTGGA	AGAAAGCATT	GCAAACCTCA	ATAGGTCGAC	CAGCGTTCCA	GAGAACCCCA	4980
AGAGTTCAAC	GTCTGCAACT	GGGTGTGCTA	CCCCGACAGA	GAAGGAGTTT	CCCAAACTC	5040
ACTCGGATGC	ATCTTCAGAA	CAGCAAAATC	GAAAAAGCCA	GACCGGCACC	AACGGAGGCA	5100
GTGTGAAATT	GATATCCACA	GACCAAGCA	CCTTTGACCT	CTTGAAGGAT	TTGGAGTTTT	5160
CCGCTGGGTC	CCCAAGTAAA	GACACAAACG	AGAGTCCCTG	GAGATCAGAT	CTGTTGATAG	5220
ATGAAAACTT	GCTTTCTCCT	TTGGCGGGAG	AAGATGATCC	ATTCCTTCTC	GAAGGGAACA	5280
CGAATGAGGA	TTGTAAGCCT	CTTATTTTAC	CGGACACTAA	ACCTAAAAT	AAGGATACTG	5340
GAGATACAAT	CTTATCAAGT	CCCAGCAGTG	TGGCACTACC	CCAAGTGAAG	ACAGAAAAAG	5400
ATGATTTTCAT	TGAACTTTGC	ACCCCCGGGG	TAATTAAGCA	AGAGAACTG	GGCCAGTTT	5460
ATTGTGAGGC	AAGCTTTTCT	GGGACAAATA	TAATTGGTAA	TAAATGTCT	GCCATTTCTG	5520
TTATCGGTGT	GAGTACCTCT	GGAGGACAGA	TGTACCACTA	TGACATGAAT	ACAGCATCCC	5580
TTTCTCAGCA	GCAGGATCAG	AAGCCTGTTT	TTAATGTCTA	TCCACCAATT	CCTGTTGGTT	5640
CTGAAAACTG	GAATAGGTGC	CAAGGCTCCG	GAGAGGACAG	CCTGACTTCC	TTGGGGGCTC	5700
TGAACCTCCC	AGGCCCGTCA	GTGTTTTCTA	ATGGGTACTC	AAGCCCTGGA	ATGAGACCAG	5760
ATGTAAGCTC	TCCTCCATCC	AGCTCGTCAG	CAGCCACGGG	ACCACCTCCC	AAGCTCTGCC	5820
TGGTGTGCTC	TCAGGATGTC	TCAGGATGTC	ATTACGGGGT	GCTGACATGT	GGAAGCTGCA	5880
AAGTATTCTT	TAAAGAGCA	GTGGAAGGAC	AGCACAATTA	CCTTTGTGCT	GGAAGAAACG	5940
ATTGCATCAT	TGATAAAATT	CGAAGGAAAA	ACTGCCACGC	ATGCCGCTAT	CGGAAATGTC	6000
TTACGGCTGG	AATGAACCTT	GAAGCTCGAA	AAACAAGAA	AAAAATCAAA	GGGATTCAGC	6060
AAGCCACTGC	AGGAGTCTCA	CAAGACACTT	CGGAAATCC	TAACAAAACA	ATAGTTCCTG	6120
CAGCATTACC	ACAGCTCACC	CCTACCTTGG	TGTCACTGCT	GGAGGTGATT	GAACCCG	6177

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	98 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GTACGTTTAA ACGCGGCGCG CCGTCGACCT GCAGAAGCTT ACTAGTGGTA CCCCATGGAG	60
ATCTGGATCC GAATTCACGC GTCTAGATT AATTAAGC	98

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	98 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGCCGCTTAA TTAATCTAGA ACGCGTGAAT TCGGATCCAG ATCTCCATGG GGTACCACTA	60
GTAAGCTTCT GCAGGTCGAC GGCGCGCCGC GTTTAAAC	98

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	51 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATCTCGGTC TCCAACAGCA ACAGCAACAG CAACAGCAAC AGGGTCTTCT G	51
----------------------------------------------------------	----

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	51 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GATCCAGAAG ACCCTGTTGC TGTGCTGTT GCTGTTGCTG TTGGAGACCG A	51
---------------------------------------------------------	----

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	42 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AATTCCCCGA GCGGCAGCT GAAATCATCA CCAATCAGAT CT

42

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	18 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TATGCCTTAC CATGTGGC

18

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	25 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TTGGTCGACA AGATCATGCA TTATC

25

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	28 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTGTGACCC GCAGTACAGA TGAAGTTG

28

## (2) INFORMATION FOR SEQ ID NO: 10:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	30 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTGGTCGACC CAGCAATAAC TTCAGACATC

30

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	29 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CGACAGATCT GGCTCCTGAG CAAAGAGAA

29

## (2) INFORMATION FOR SEQ ID NO: 12:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	24 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CCAGGGATCC TCTCCTTGCT GCAA

24

## (2) INFORMATION FOR SEQ ID NO: 13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	33 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCTAGTCGAC GATGGCTCCT GAGCAAAGAG AAG

33

## (2) INFORMATION FOR SEQ ID NO: 14:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	27 base pairs
(B) TYPE:	nucleic acid
(C) STRANDEDNESS:	single
(D) TOPOLOGY:	linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CCAGGGATCC TATCCTTGCT GCAACAG

27

08959013-103397